

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of

Tadashige KADOI

Serial No.: 10/627,740

Group Art Unit: Not Yet Assigned

Filing Date: July 28, 2003

Examiner: Unknown

For: MULTI-PROCESSOR SYSTEM

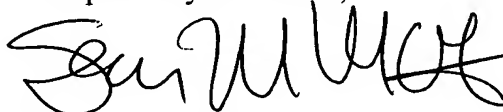
Honorable Commissioner of Patents
Alexandria, VA 22313-1450

SUBMISSION OF PRIORITY DOCUMENT

Sir:

Submitted herewith is a certified copy of Japanese Application Number 2002-220125
filed on July 29, 2002, upon which application the claim for priority is based.

Respectfully submitted,



Sean M. McGinn, Esq.
Registration No. 34,386

Date: 10/8/03

McGinn & Gibb, PLLC
Intellectual Property Law
8321 Courthouse Road, Suite 200
Vienna, VA 22182-3817
(703) 761-4100
Customer No. 21254

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application: 2002年 7月29日

出 願 番 号

Application Number: 特願2002-220125

[ST.10/C]:

[JP2002-220125]

出 願 人

Applicant(s): 日本電気株式会社

2003年 5月 6日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田信一郎

出証番号 出証特2003-3032934

【書類名】 特許願

【整理番号】 66206602

【提出日】 平成14年 7月29日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 15/177 678

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 門井 忠茂

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100083839

 【弁理士】

 【氏名又は名称】 石川 泰男

 【電話番号】 03-5443-8461

【選任した代理人】

 【識別番号】 100109139

 【弁理士】

 【氏名又は名称】 今井 孝弘

 【電話番号】 03-5443-8461

【選任した代理人】

 【識別番号】 100111062

 【弁理士】

 【氏名又は名称】 樋口 直篤

 【電話番号】 03-5443-8461

【手数料の表示】

 【予納台帳番号】 007191

 【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0116381

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 マルチプロセッサシステムの障害処理方法、マルチプロセッサシステム及びノード

【特許請求の範囲】

【請求項1】 複数のノードと前記複数のノードを制御するサービスプロセッサとを有するノード群が複数集まって構成され、前記複数のノード群における複数の前記ノードは相互接続網により相互に接続され、複数の前記サービスプロセッサは該サービスプロセッサを制御するサービスプロセッサマネージャに接続され、複数の前記ノード群から選択された複数の前記ノードでパーティションを構成したマルチプロセッサシステムにおける障害処理方法であって、

いずれかの前記ノードで障害が発生した場合に、その障害ノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該サービスプロセッサは、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する工程と、

前記障害ノードは、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達する工程と、

前記障害ノードと同一のパーティションに属する他のノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該サービスプロセッサは、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する工程と、

前記サービスプロセッサマネージャは、前記障害情報に基づいて前記それぞれのサービスプロセッサに障害処理を指示し、この指示を受けた前記サービスプロセッサは、該指示に従って障害処理をおこなう工程と、を含むことを特徴とするマルチプロセッサシステムの障害処理方法。

【請求項2】 請求項1に記載のマルチプロセッサシステムにおける障害処理方法において、

前記障害ノードは、前記相互接続網を用いて障害通知パケットを他のノードに送信して、前記障害ノードと同一のパーティションに属する他のノードに前記他のノードに障害情報を伝達することを特徴とするマルチプロセッサシステムの障

害処理方法。

【請求項 3】 請求項 1 に記載のマルチプロセッサシステムにおける障害処理方法において、

前記障害ノードは、前記相互通信網を用いて障害通知パケットを前記障害ノードと同一のパーティションに属する他のノードを宛先として送信して、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達することを特徴とするマルチプロセッサシステムの障害処理方法。

【請求項 4】 請求項 1 に記載のマルチプロセッサシステムにおける障害処理方法において、

前記障害ノードは、前記相互通信網を用いて障害通知パケットをブロードキャストで送信し、障害通知パケットを受信したノードが前記障害ノードと同一パーティションに属する場合のみ前記障害情報を該ノードに取り込むことで前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達することを特徴とするマルチプロセッサシステムの障害処理方法。

【請求項 5】 請求項 2 乃至 4 のいずれかに記載のマルチプロセッサシステムにおける障害処理方法において、

前記障害ノードは、前記相互接続網を用いて障害通知パケットを送信する際に、通常のトランザクションで用いるチャンネルとは異なるチャンネルを用いて障害通知パケットを送信して、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達することを特徴とするマルチプロセッサシステムの障害処理方法。

【請求項 6】 複数のノードが該ノードを制御するサービスプロセッサに接続されてなるノード群が複数集まって構成され、複数の前記ノードは相互接続網により相互に接続され、前記サービスプロセッサおよびサービスプロセッサを制御するサービスプロセッサマネージャが専用線で接続され、複数の前記ノード群から選択された複数の前記ノードでパーティションを構成したマルチプロセッサシステムであって、

いずれかの前記ノードで障害が発生した場合に、その障害ノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該

サービスプロセッサは、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する障害ノード通知手段と、

前記障害ノードは、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達するパーティション内通知手段と、

前記障害ノードと同一のパーティションに属する他のノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該サービスプロセッサは、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する他ノード通知手段と、

前記サービスプロセッサマネージャは、前記障害情報に基づいて前記サービスプロセッサに障害処理を指示し、この指示を受けた前記サービスプロセッサは、該指示に従って障害処理をおこなう障害処理手段と、を備えたことを特徴とするマルチプロセッサシステム。

【請求項 7】 複数のノードが該ノードを制御するサービスプロセッサに接続されてなるノード群が複数集まって構成され、複数の前記ノードは相互接続網により相互に接続され、前記サービスプロセッサおよびサービスプロセッサを制御するサービスプロセッサマネージャが専用線で接続され、複数の前記ノード群から選択された複数の前記ノードでパーティションを構成したマルチプロセッサシステムに含まれるノードであって、

該ノードに障害が発生した場合に、該ノードと同一のパーティションに属する他のノードに障害情報を通知するとともに、該ノードを制御する前記サービスプロセッサに障害発生を通知することを特徴とするノード。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、マルチプロセッサシステムの障害処理方法、マルチプロセッサシステム及びノードに関する。

【0002】

【従来の技術】

近年、マルチプロセッサシステム、特にWindows（登録商標）やUNIX

X（登録商標）等のオープンシステムにおいて、プラットフォームがOS（Operating System）、ドライバ、アプリケーションと協力して、システムの構成制御、電源制御、エラーロギング、障害処理等をおこなうRAS（Remote Access Service）機能を強化する傾向にある。

【0003】

また、システム基盤であるプラットフォームは、ユーザニーズの多様化に対応するため大規模化しており、特に、複数のプロセッサを有するマルチプロセッサシステムを独立して運用可能なシステム（パーティション）に分割して、複数のOSを運用可能とする機能が要求されている。

【0004】

このような状況から今後、大規模プラットフォームを複数のパーティションに分割し、それぞれのパーティションの負荷状態に応じて柔軟にリソースの追加又は削除をおこなう機能、予備リソースを用意しておき障害発生時に即座に自動リプレースする機能等への要求に応えることができ、また、システム統合によるコスト削減を実現できる、統合プラットフォームへのニーズが高まることが予想される。

【0005】

ところで、マルチプロセッサシステムを用いた場合は、システム障害に対する復旧処理を的確におこなうことが重要となる。例えば、複数のノードを1つのサービスプロセッサで制御しておこなうマルチプロセッサシステムに対しての障害処理方法が開示されている（特開2001-134546公報）。

【0006】

【発明が解決しようとする課題】

しかし、複数のノードをサービスプロセッサで制御してなるノード群が複数集まって構成されたシステムで、異なるノード群に属する複数のノードを選択して独立のシステムとして運用するマルチプロセッサシステムにおいては、障害が複数のノード群に渡って発生する可能性があり、上記の障害処理手段では的確に障害処理をおこなうことができない。

【0007】

本発明はこのような問題点に鑑みて提案されたものであり、複数のノードが該ノードを制御するサービスプロセッサに接続されてなるノード群が複数集まって形成された大規模プラットフォームを用いたマルチプロセッサシステムにおいて、障害処理を迅速かつ的確におこなうことができる障害処理方法、マルチプロセッサシステム及びノードを提供することを目的とする。

【 0 0 0 8 】

【課題を解決するための手段】

上記課題を解決するため、請求項 1 に記載の発明は、複数のノードと前記複数のノードを制御するサービスプロセッサとを有するノード群が複数集まって構成され、前記複数のノード群からなる複数の前記ノードは相互接続網により相互に接続され、複数の前記サービスプロセッサは該サービスプロセッサを制御するサービスプロセッサマネージャに接続され、複数の前記ノード群から選択された複数の前記ノードでパーティションを構成したマルチプロセッサシステムにおける障害処理方法であって、いずれかの前記ノードで障害が発生した場合に、その障害ノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該サービスプロセッサは、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する工程と、前記障害ノードは、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達する工程と、前記障害ノードと同一のパーティションに属する他のノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該サービスプロセッサは、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する工程と、前記サービスプロセッサマネージャは、前記障害情報に基づいて前記それぞれのサービスプロセッサに障害処理を指示し、この指示を受けた前記サービスプロセッサは、該指示に従って障害処理をおこなう工程と、を含むことを特徴とする。

【 0 0 0 9 】

請求項 1 に記載の発明によれば、障害ノードと同一パーティション内のノードから各ノードの障害情報を入手して、それに基づいて障害処理をおこなうので、被疑部の特定を迅速かつ的確におこなえ、ひいては障害処理を迅速かつ的確にお

こなうことができる。なお、パーティションとはシステムの一部であり、独立して運用可能ものをいう。また、障害情報とは障害がある旨の情報、または、エラーログなどの障害内容を示す情報をいう。

【 0 0 1 0 】

請求項 2 に記載の発明は、請求項 1 に記載のマルチプロセッサシステムにおける障害処理方法において、前記障害ノードは、前記相互接続網を用いて障害通知パケットを他のノードに送信して、前記障害ノードと同一のパーティションに属する他のノードに前記他のノードに障害情報を伝達することを特徴とする。

【 0 0 1 1 】

請求項 2 に記載の発明によれば、障害情報を障害パケットの送信でおこなうので、通常のトランザクションにおけるパケット送信と同じ環境を用いて障害処理が可能となる。

【 0 0 1 2 】

請求項 3 に記載の発明は、請求項 1 に記載のマルチプロセッサシステムにおける障害処理方法において、前記障害ノードは、前記相互通信網を用いて障害通知パケットを前記障害ノードと同一のパーティションに属する他のノードを宛先として送信して、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達することを特徴とする。

【 0 0 1 3 】

請求項 3 に記載の発明によれば、障害パケットを前記障害ノードと同一のパーティションに属する他のノードを宛先として送信するので、不要なパケットを相互接続網に送信せずに済み、効率的な障害処理が可能となる。

【 0 0 1 4 】

請求項 4 に記載の発明は、請求項 1 に記載のマルチプロセッサシステムにおける障害処理方法において、前記障害ノードは、前記相互通信網を用いて障害通知パケットをブロードキャストで送信し、障害通知パケットを受信したノードが前記障害ノードと同一パーティションに属する場合のみ前記障害情報を該ノードに取り込むことで前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達することを特徴とする。

【 0 0 1 5 】

請求項 4 に記載の発明によれば、障害ノードで障害通知パケットの宛先を特定しないので障害ノードでの処理負荷を低減でき、障害ノードの処理負荷が大きいときに有利である。

【 0 0 1 6 】

請求項 5 に記載の発明は、請求項 2 乃至 4 のいずれかに記載のマルチプロセッサシステムにおける障害処理方法において、

前記障害ノードは、前記相互接続網を用いて障害通知パケットを送信する際に、通常のトランザクションで用いるチャンネルとは異なるチャンネルを用いて障害通知パケットを送信して、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達することを特徴とする。

【 0 0 1 7 】

請求項 5 に記載の発明によれば、障害の影響等で通常のトランザクションに用いられるチャンネルが渋滞していたり、性能が低下していたりしても迅速、かつ、確実に障害通知パケットを送信することが可能となる。

【 0 0 1 8 】

請求項 6 に記載の発明は、複数のノードが該ノードを制御するサービスプロセッサに接続されてなるノード群が複数集まって構成され、複数の前記ノードは相互接続網により相互に接続され、前記サービスプロセッサおよびサービスプロセッサを制御するサービスプロセッサマネージャが専用線で接続され、複数の前記ノード群から選択された複数の前記ノードでパーティションを構成したマルチプロセッサシステムであって、いずれかの前記ノードで障害が発生した場合に、その障害ノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該サービスプロセッサは、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する障害ノード通知手段と、前記障害ノードは、前記障害ノードと同一のパーティションに属する他のノードに障害情報を伝達するパーティション内通知手段と、前記障害ノードと同一のパーティションに属する他のノードは、該ノードを制御する前記サービスプロセッサに障害の発生を通知し、この通知により、該サービスプロセッサは

、該ノードから障害情報を取得して、該障害情報を前記サービスプロセッサマネージャに送信する他ノード通知手段と、前記サービスプロセッサマネージャは、前記障害情報に基づいて前記サービスプロセッサに障害処理を指示し、この指示を受けた前記サービスプロセッサは、該指示に従って障害処理をおこなう障害処理手段と、を備えたことを特徴とする。

【 0 0 1 9 】

請求項 7 に記載の発明は、複数のノードが該ノードを制御するサービスプロセッサに接続されてなるノード群が複数集まって構成され、複数の前記ノードは相互接続網により相互に接続され、前記サービスプロセッサおよびサービスプロセッサを制御するサービスプロセッサマネージャが専用線で接続され、複数の前記ノード群から選択された複数の前記ノードでパーティションを構成したマルチプロセッサシステムに含まれるノードであって、該ノードに障害が発生した場合に、該ノードと同一のパーティションに属する他のノードに障害情報を通知するとともに、該ノードを制御する前記サービスプロセッサに障害発生を通知することを特徴とする。

【 0 0 2 0 】

請求項 6 または 7 に記載の発明によれば、何れかのノードで障害が発生した場合、相互接続網を用いて同一パーティションに属する他のノードに障害情報を通知して、該障害情報に基づいて障害処理をおこなえるので、また、サービスプロセッサマネージャが各サービスプロセッサを制御して障害処理をおこなえるので迅速かつ的確な障害処理を可能となる。

【 0 0 2 1 】

【発明の実施の形態】

以下、本発明の好適な実施の形態について、添付図面を参照して説明する。

《システム構成》

図 1 は、本発明の実施形態にかかるマルチプロセッサシステムの一例を示すシステム構成図である。

【 0 0 2 2 】

図 1 において、1 はマルチプロセッサシステムを、2 はノード群を、3 はノード

ドを、4はサービスプロセッサを、5はプロセッサを、6メモリを、7はノースブリッジを、8はI/Oホストブリッジを、9はクロスバーを、10は相互接続網を、11はサービスプロセッサマネージャを、12は専用線を示している。

【0023】

マルチプロセッサシステム1はノード群A～Dの4つのノード群2で構成されている。マルチプロセッサシステム1を構成するノード群2の数は本実施形態では4つであるが、これに限定されず、本発明は2つ以上のノード群2で構成される場合に適用が可能である。各ノード群2はそれぞれ遠隔していても良いし、隣接していても良いが、遠隔して設置されている場合には、遠隔して設置されたノード3を用いてマルチプロセッサシステム1を構成できるという特徴を有する。

【0024】

ノード群AはノードAa～Ahの8つのノード3およびノードAa～Ahを管轄するサービスプロセッサA（図において、SPA）で構成されている。同様にノード群B～CはノードBa～Bh，ノードCa～Ch，ノードDa～Dhとそれぞれを管轄するサービスプロセッサB～D（図において、SPB，SPC，SPD）で構成されている。1つのノード群2が有するノード2の数は本実施形態では8つであるが、これに限定されず、本発明は2つ以上のノード3を有するノード群2で構成されるマルチプロセッサシステム1に対して適用が可能である。また、本実施形態において、ノード3は、P/Mノード3aとI/Oノード3bの2種類のノードを有するが、特にこれに限定されることなく、例えばP/Mノード3aのみでもよいし、その他の種類のノード3で構成しても良い。

【0025】

P/Mノード3aは主に演算処理をおこなうノードで、少なくともプロセッサ5とメモリ6とを備えている。また、プロセッサ5とノード3内の各パーツ間を結ぶバスを接続するノースブリッジ7を備えている。なお、本実施形態ではプロセッサ5の数は2で、メモリ6の数は1であるが、プロセッサ5、メモリ6などの数は特にこれに限定されるものではない。また、I/Oノード3bは主に入出力をおこなうノードで、I/Oホストブリッジ8と配下にI/Oデバイス（図示せず）を備えている。

【 0 0 2 6 】

サービスプロセッサ 4 は、ノード群 2 内のノード 3 を接続しており、これらノード 3 のマネージメントをおこなっている。

【 0 0 2 7 】

ノード群 A ～ D に属する計 3 2 個のノード 3 は、それぞれクロスバー 9 を介して相互接続網 1 0 によって相互に連結されている。なお、クロスバー 9 は、複数のプロセッサやメモリの間でデータをやり取りする際に、経路を動的に選択する機能を備えた装置である。

【 0 0 2 8 】

サービスプロセッサマネージャ 1 1 は、各サービスプロセッサ 4 を専用線 1 2 で接続しており、これらサービスプロセッサ 4 のマネージメントをおこなっている。

【 0 0 2 9 】

このようなハードウェアプラットフォームにおいて、複数のノード群 2 から複数のノード 3 を選択して独立して運用可能なシステムであるパーティションを形成している。具体的には、ノード群 A に属するノード A e ～ A h，ノード群 B に属するノード B a ～ B d，ノード群 C に属するノード C a ～ C h の全 1 2 個のノードを選択してパーティション 1 3 を形成している。なお、本実施形態では 1 つのパーティションのみを定義したが、当然に複数のパーティションを形成してもよい。

【 0 0 3 0 】

《障害処理方法》

次に、以上のマルチプロセッサシステムにおける障害処理方法の一例について、図 2 ～ 9 を参照して説明する。

【 0 0 3 1 】

図 2 ～ 7 は各障害処理工程の説明図、図 8 は障害処理のフロー図、図 9 は障害通知パケットの説明図である。以下、図 8 のフロー図中のステップ 1 ～ ステップ 6 のそれぞれのステップに分けて説明する。なお、ステップ 1 ～ ステップ 6 は、それぞれ図 2 ～ 図 7 に対応する。

【0032】

<ステップ1>

障害ノードが、障害ノードを管轄するサービスプロセッサに、障害を通知するステップである。

【0033】

図2に示すように、ノード群Aに属するノードAeのI/Oホストブリッジ8でデータ転送中に重障害、例えば内部ECCエラー等が発生したとして、以下説明する。ノードAeは、障害を検知したら、まず内部のトレース情報、レジスタ情報などのエラーログ情報（障害情報）をホールドする。次に、ノードAeは、続いておこなう予定であったデータ転送を中止するとともに、ノードAeを管轄するサービスプロセッサA（SPA）に障害発生を送信し、SPAはノードAeの障害情報を取得する。障害情報を受け取ったSPAは、害障害の処理がノードAeの切り離しで済むか、または、パーティションリセットが必要かを判断する。SPAは、パーティションリセットが必要と判断した場合は、即座にノードAeをリセットし、サービスプロセッサマネージャ11に障害情報を送信してパーティションリセットのリクエストをおこなう。

【0034】

<ステップ2>

障害ノードが、同一パーティションに属する他のノードに、障害を通知するステップである。ステップ2は、ステップ1と並行しておこなわれる。

【0035】

障害ノードであるノードAeは、他のノードに障害発生を通知するため、障害通知パッケージを作成する。図9に示すように、障害パッケージは、エラーコード、宛先ノードコード、送信元ノードコード、重障害フラグ、エラー情報で構成される。エラーコードは、障害通知パッケージであることを示す符号である。宛先ノードコードは、障害通知パッケージの宛先を示す符号で、ブロードキャストアドレスが指定されている。送信元ノードコードは、障害通知パッケージの送信元を示す符号で、障害ノードであるノードAeのアドレスが指定されている。重障害フラグは、障害が重度か否かを示す符号で、ノードAeでは重障害が発生しているので、

フラグは立っている。エラー情報は、ノードA eでの障害情報を示している。

【0036】

図3に示すように、障害通知パケットは、相互接続網を介して他のすべてのノードにブロードキャストで送信される。このとき通常のトランザクションに用いられるチャンネルとは異なるチャンネルを用いて送信される。したがって、通常のトランザクションに用いられるチャンネルが障害により渋滞していたり、性能が低下していたりしても迅速、確実に障害通知パケットを送信するためである。

【0037】

障害通知パケットを受信したノードは、ノースブリッジ7またはI/Oホストブリッジ8に設定されているパーティション情報で、自己ノードが障害パケットの送信元ノードであるノードA eと同一パーティションに属しているかを判断し、属していると判断した場合のみ障害通知パケットを自己ノードのエラーログ情報の一部として保存する。したがって、障害ノードA eと同一パーティションに属するノードA f～A h、ノードB a～B d、ノードC a～C hは、それぞれのノードのエラーログ情報の一部として、障害通知パケットを保存する。

【0038】

<ステップ3>

障害ノードと同一のパーティションに属するノードが、該ノードを管轄するサービスプロセッサに障害を通知する工程である。

【0039】

図4に示すように、障害ノードA eと同一パーティションに属し、障害通知パケットをエラーログ情報の一部として保存した、ノードA f～A h、ノードB a～B d、ノードC a～C hは、障害通知パケットの内容を確認し、重障害フラグが立っていることを認識すると、自己のエラーログ情報をホールドするとともに、管轄のサービスプロセッサであるSPA、SPB、SPCにそれぞれ障害発生を通知する。

【0040】

通知を受けたSPA、SPB、SPCは、管轄下のノードのエラーログ情報を取得し、障害通知パケットの内容に基づいて、関連するノードをリセットする。

【0041】

<ステップ4>

障害ノードと同一パーティションに属するノードを管轄するサービスプロセッサが、サービスプロセッサマネージャに障害を通知する工程である。

【0042】

図5に示すように、障害の通知を受けたSPA、SPB、SPCは、管轄下のノード（それぞれ、ノードAf～Ah、ノードBa～Bd、ノードCa～Ch）の障害情報をサービスプロセッサマネージャ11に送信する。

【0043】

<ステップ5>

サービスプロセッサマネージャが、障害被疑部の特定をおこなう工程である。

【0044】

図6に示すように、関連するそれぞれのノード（ノードAe～Ah、ノードBa～Bd、ノードCa～Ch）の障害情報を取得したサービスプロセッサマネージャ11は、障害通知パケット情報が一致していることにより、それぞれが同一原因であることを認識する。そして、サービスプロセッサマネージャ11が保有する故障解析辞書に基づいて、障害被疑部の特定をおこなう。また、障害被疑部の特定と並行して、障害情報を論理パーミッション情報、物理ロケーション情報などのシステム構成情報と組合せてログ情報管理をおこなう。

【0045】

<ステップ6>

障害回復処理をおこなう工程である。

【0046】

図7に示すように、パーティションリセットが必要であると判断された場合、サービスプロセッサマネージャ11は、SPA、SPB、SPCにパーティションのリセットの指示をおこなう。SPA、SPB、SPCは、相互に同期を取りながらリセットを実現する。なお、パーティションリセットは、サービスプロセッサマネージャ11の完全従属制御により、サービスプロセッサマネージャ11からの一連のリセット手順に基づいておこなっても良い。

【0047】

この他にも、障害回復処理の方法は様々であり、障害の状況に応じて回復処理がなされる。例えば、ミッションクリティカルな運用下では、サービスプロセッサマネージャ11と保守センター（図示せず）を回線で接続して、保守センターが障害処理をサポートすることが通常なされる。保守センターの保守員は、サービスプロセッサマネージャ11からリモートで障害情報を採取し、これらの障害情報に基づいて、最小ロスで、迅速かつ正確な部品交換などをおこなうことが可能である。

【0048】

また、障害ノードの代替リソースがシステム内に用意されている場合は、パーティションリセット時に、障害ノードの変えて代替リソースを組み込んでリセットをおこなえば、リソースの減少によるシステムの高負荷運用を防止することができる。

【0049】

さらに、OSのRAS機能が強化されている場合は、障害ノードの切り離して障害回復がおこなえるときに、パーティションリセットをおこなわずに、障害ノードを代替リソースと置き換えて処理することで、システムのロバストネスを実現することができる。

【0050】

以上説明したように、本実施形態によると、複数のノードが該ノードを制御するサービスプロセッサに接続されてなるノード群が複数集まって形成された大規模プラットフォームを用いたマルチプロセッサシステムにおいて障害が発生した場合に、障害箇所を迅速かつ的確に特定することができ、他のパーティションに障害を波及させることなく、また、障害が発生したパーティションの障害処理を迅速かつ的確におこなうことができる。さらに、これによりミッションクリティカルな分野への、大規模なオープンシステムの適用を可能にする。

【0051】

《他の実施形態》

上記実施形態のステップ2において、障害ノードは、障害ノードと同一パーテ

ィションに属するノードにのみ、障害通知パケットを送信しても良い。この場合、障害通知パケットの宛先ノードコードには、障害ノードと同一パーティションに属するノードのアドレスが定義される。このようにして、障害ノードと同一パーティションに属するノードにのみに障害通知パケットを送信した場合は、受信ノードでパーティションの確認をする必要がなく、即座に続く処理をおこなうことができる。

【 0 0 5 2 】

また、同ステップ2において、障害ノードは、障害ノードと同一のノード群に属するノードに対しておこなう障害通知を、障害通知パケットの送信によりおこなうのに変えて、管轄するサービスプロセッサでの折り返しや、クロスバーでの折り返しによっておこなっても良い。

【 0 0 5 3 】

【発明の効果】

以上の説明した通り、本発明によれば、障害ノードと同一パーティション内のノードから各ノードの障害情報を入手して、それに基づいて障害処理をおこなうので、被疑部の特定を迅速かつ的確におこなえ、ひいては障害処理を迅速かつ的確におこなうことができる。

【図面の簡単な説明】

【図1】 本発明の実施形態にかかるマルチプロセッサシステムの一例を示すシステム構成図である。

【図2】 本発明の実施形態にかかるマルチプロセッサシステムにおける障害処理の一工程を示す図である。

【図3】 本発明の実施形態にかかるマルチプロセッサシステムにおける障害処理の一工程を示す図である。

【図4】 本発明の実施形態にかかるマルチプロセッサシステムにおける障害処理の一工程を示す図である。

【図5】 本発明の実施形態にかかるマルチプロセッサシステムにおける障害処理の一工程を示す図である。

【図6】 本発明の実施形態にかかるマルチプロセッサシステムにおける障害処

理の一工程を示す図である。

【図 7】 本発明の実施形態にかかるマルチプロセッサシステムにおける障害処理の一工程を示す図である。

【図 8】 本発明の実施形態にかかるマルチプロセッサシステムにおける障害処理方法の一例の処理フローを示す図である。

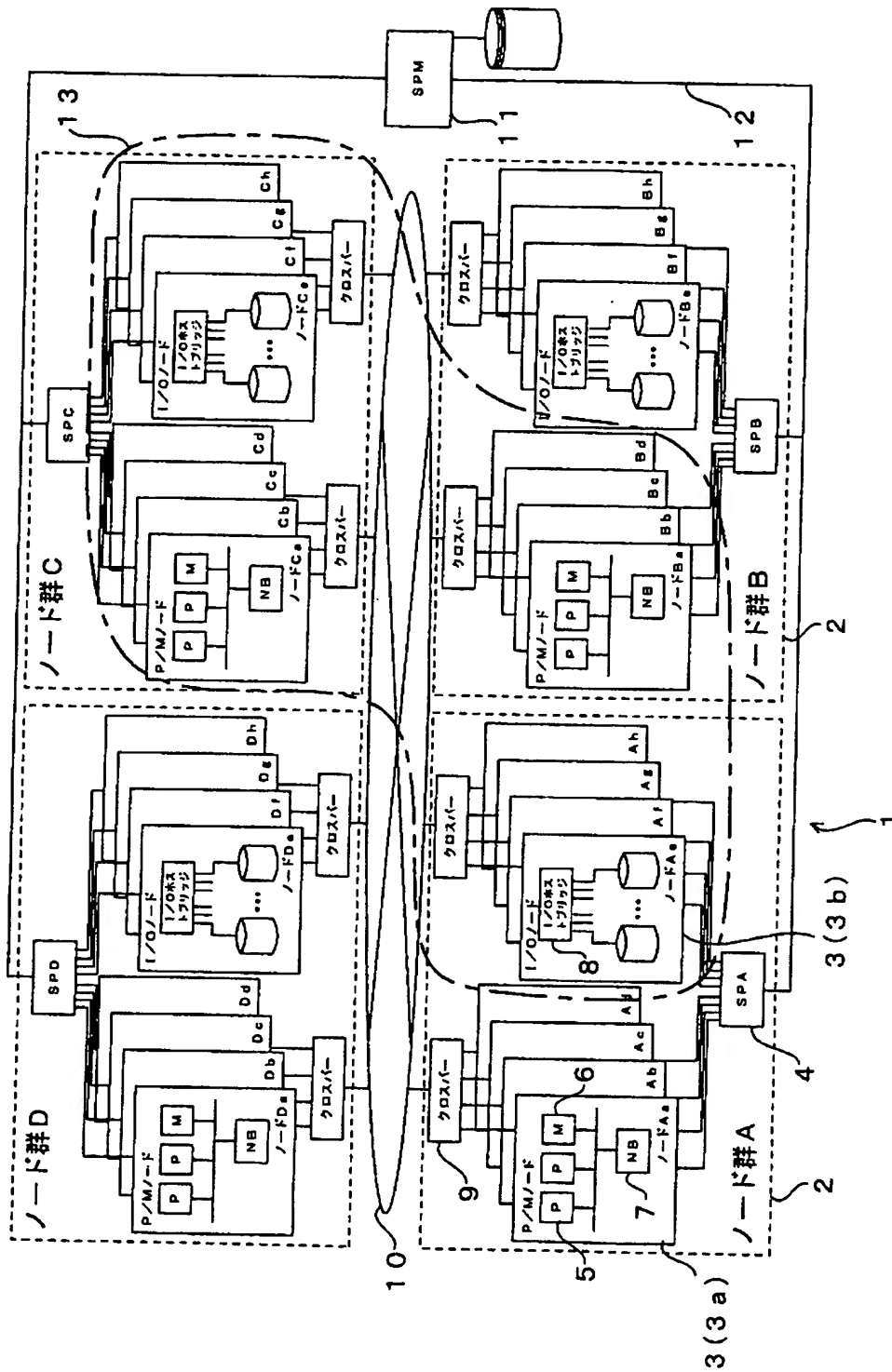
【図 9】 障害通知パケットの一例を示す図である。

【符号の説明】

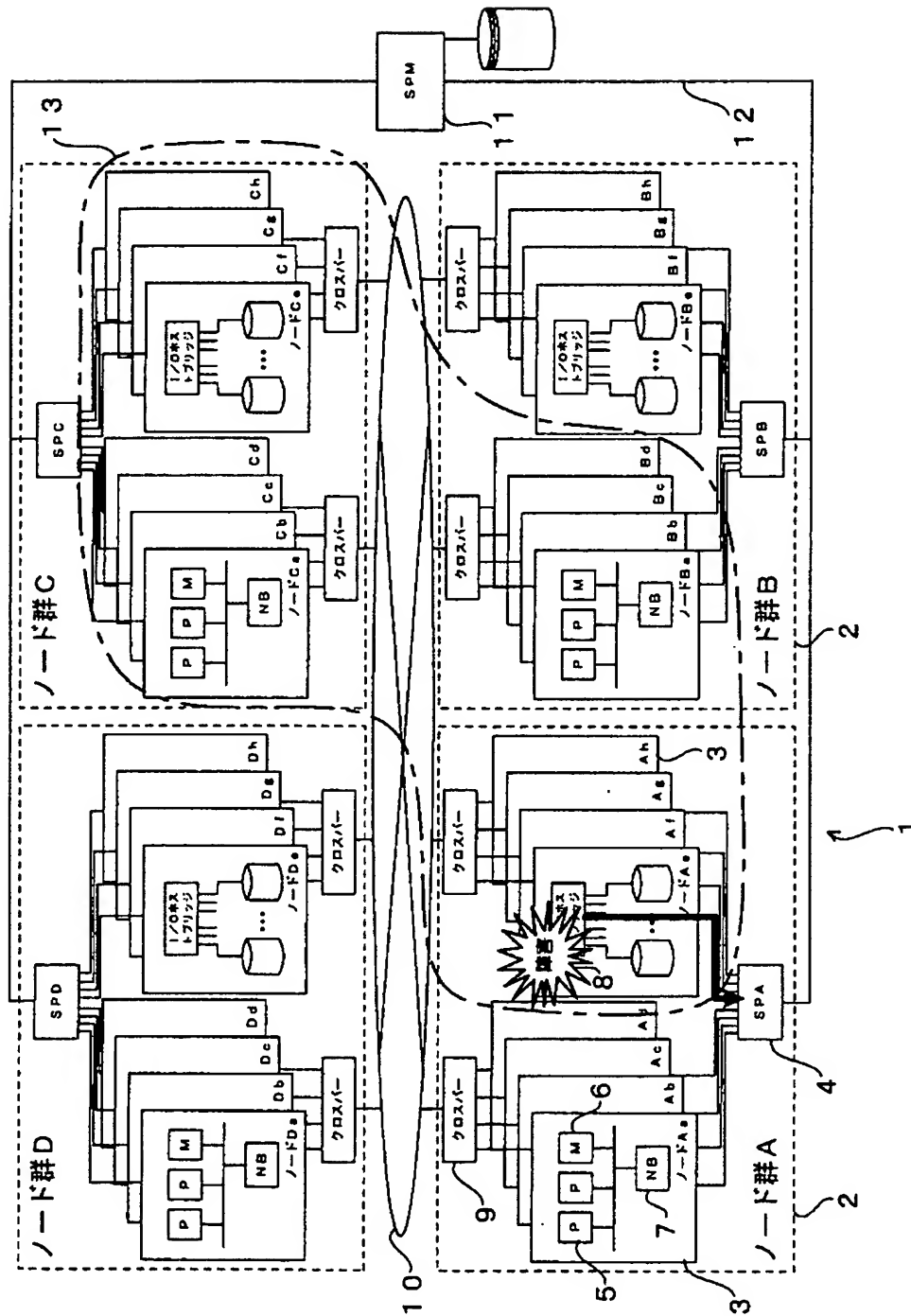
- 1 マルチプロセッサシステム
- 2 ノード群
- 3 ノード
- 4 サービスプロセッサ
- 5 プロセッサ
- 6 メモリ
- 7 ノースブリッジ
- 8 I/Oホストブリッジ
- 9 クロスバー
- 10 相互接続網
- 11 サービスプロセッサマネージャ
- 12 専用線

【書類名】 図面

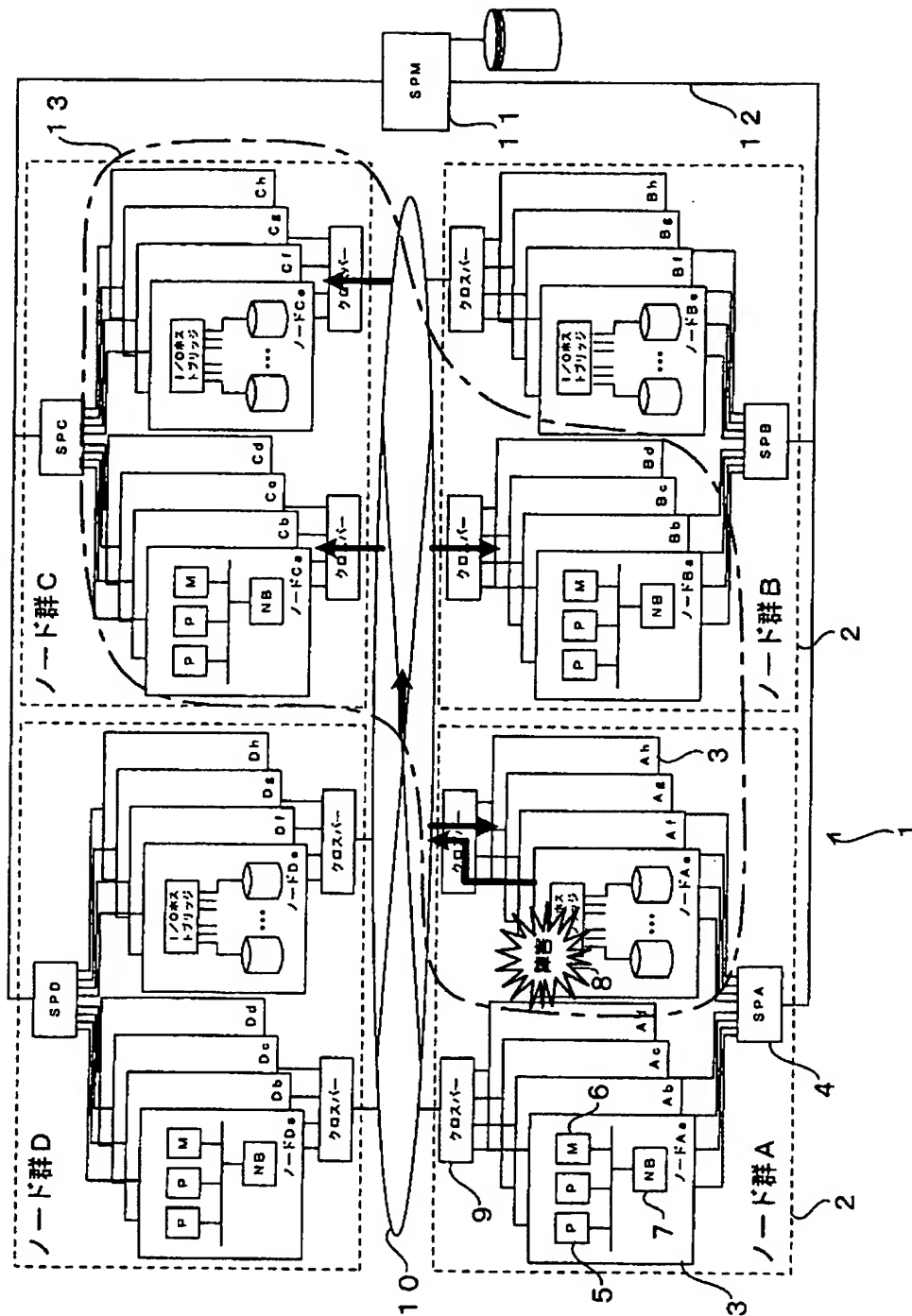
【図 1】



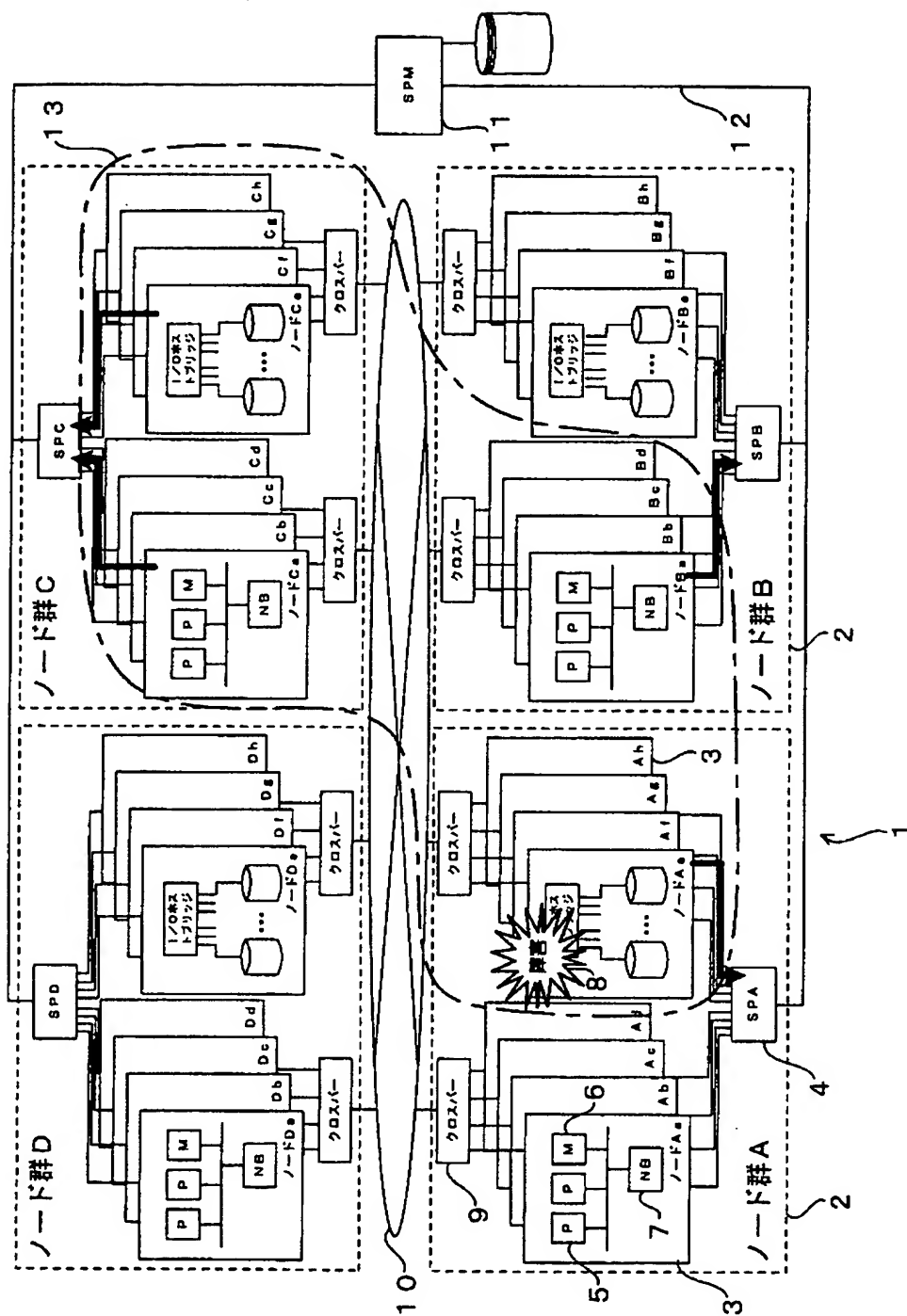
【図2】



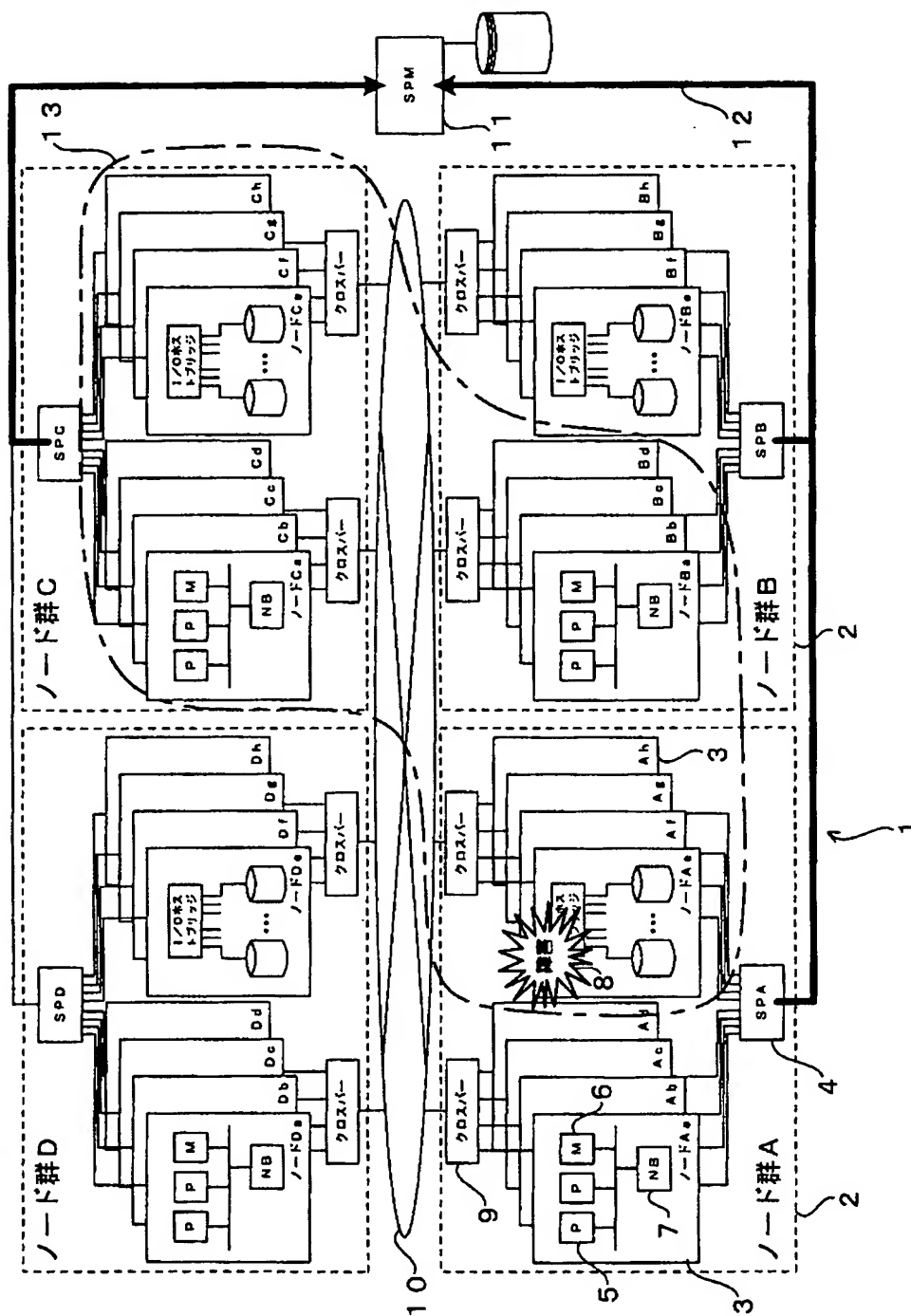
【図 3】



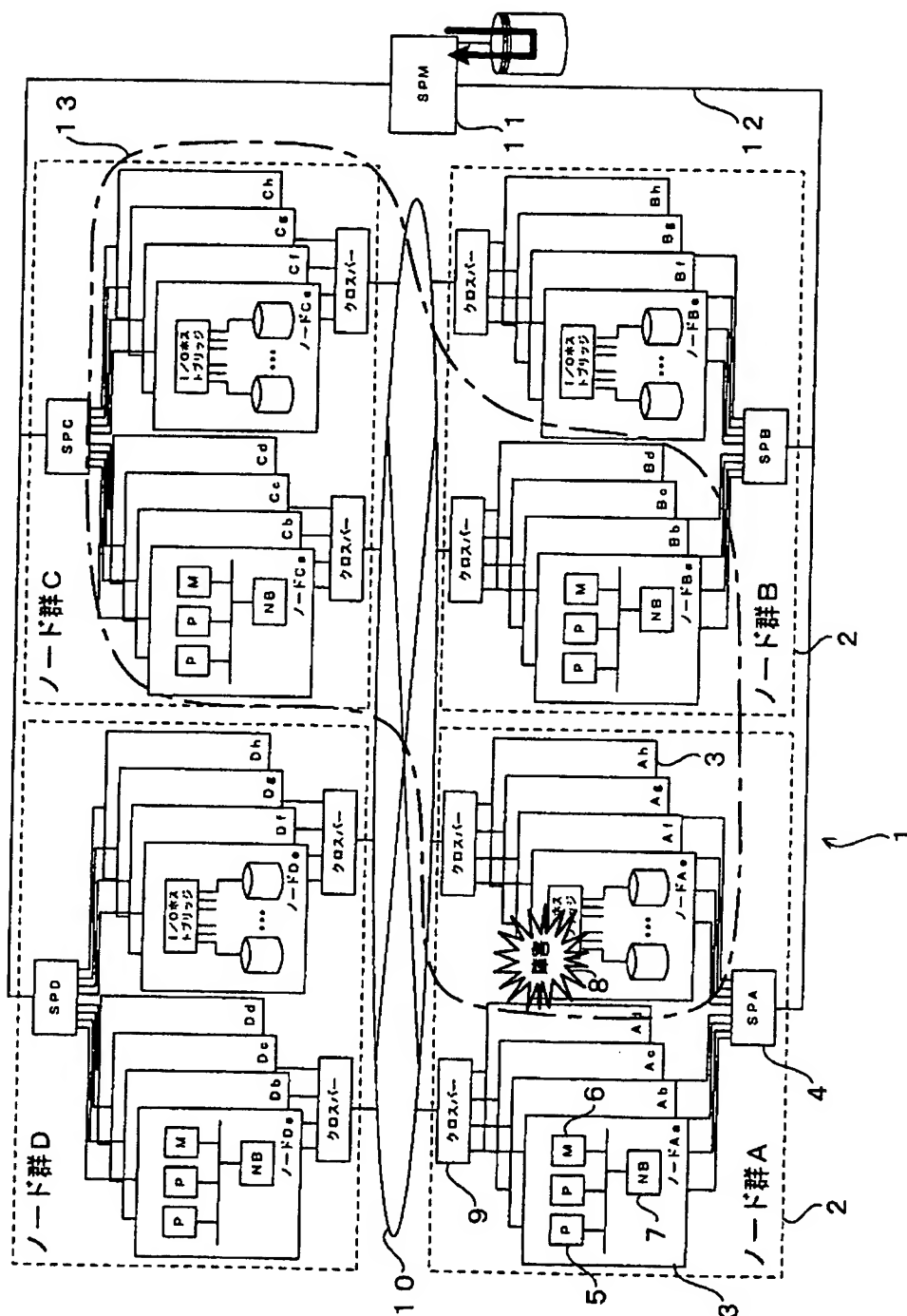
【図4】



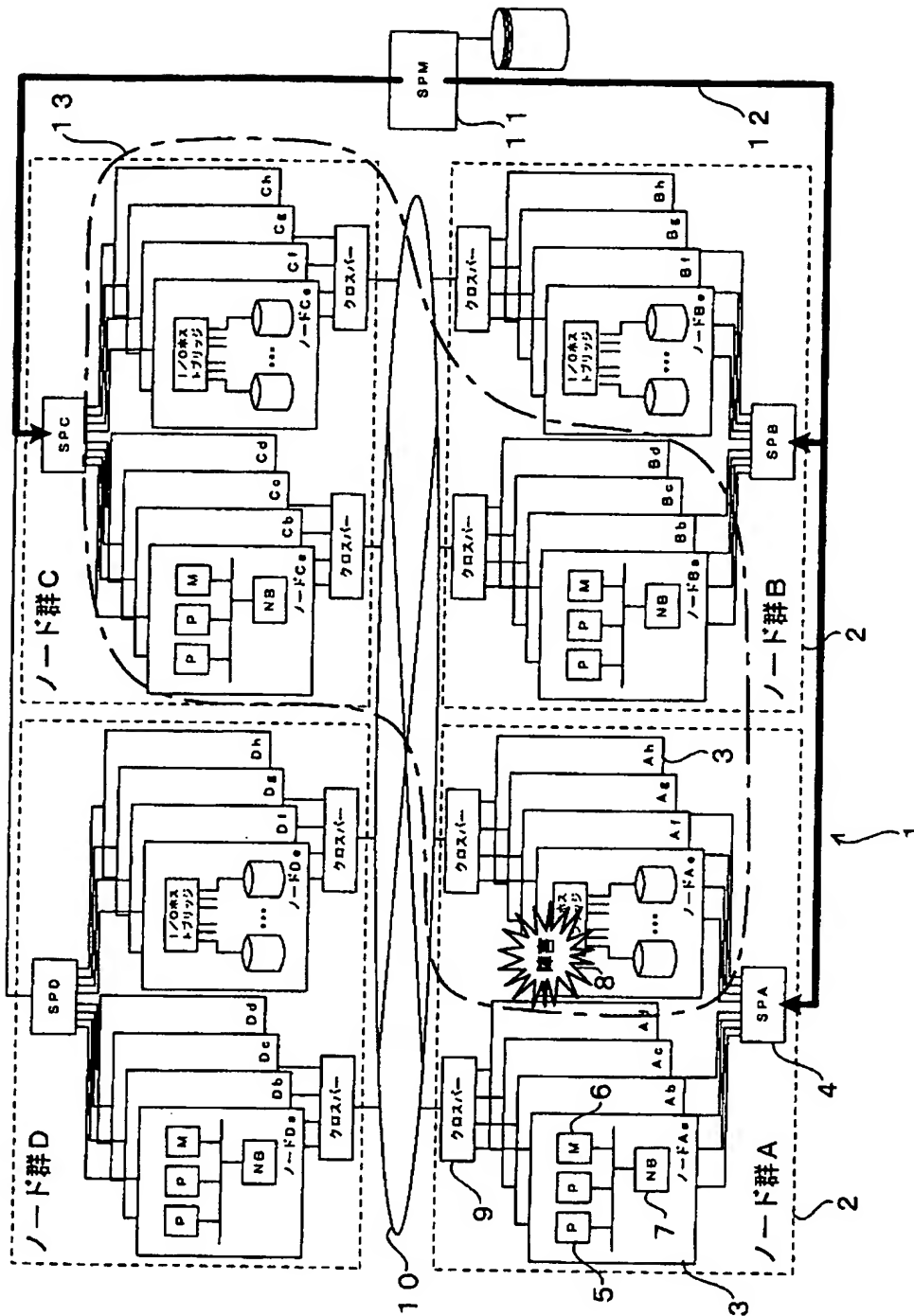
【図5】



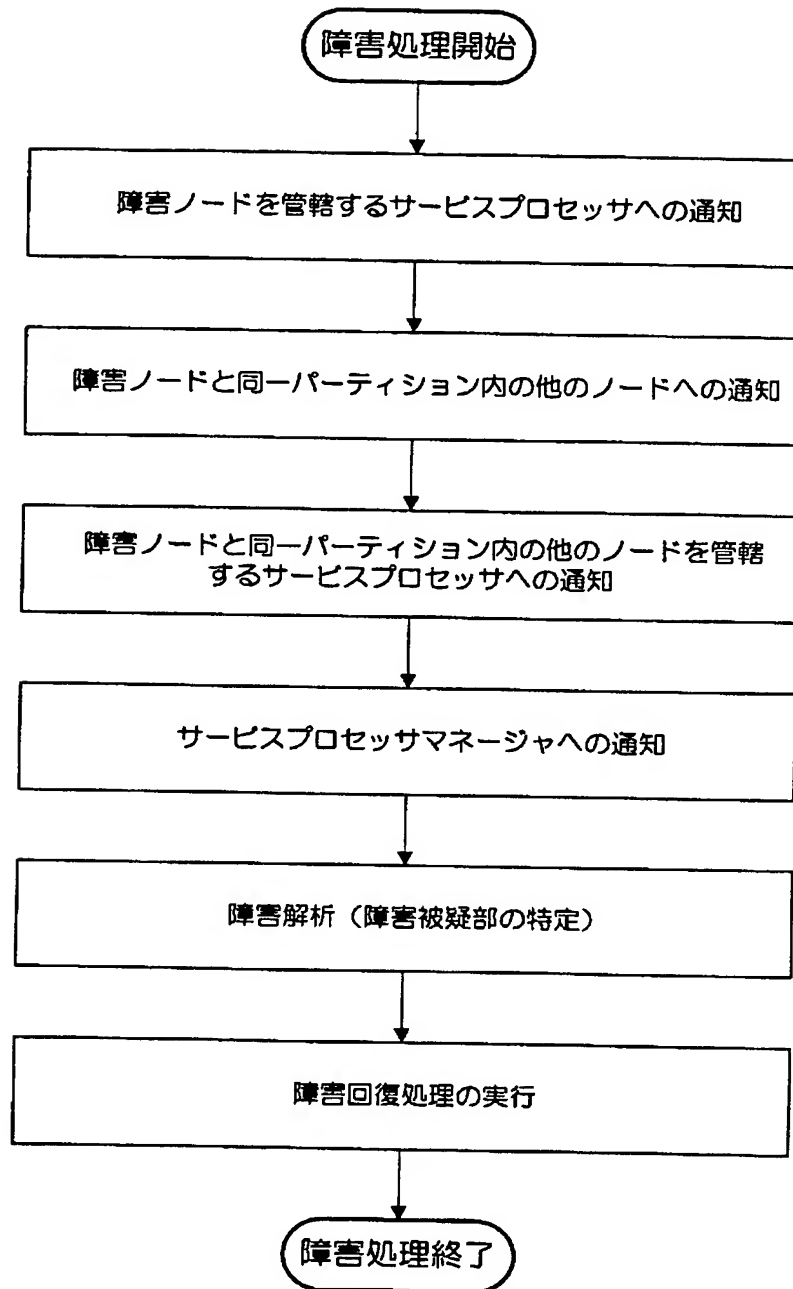
【図6】



【図7】



【図 8】



【図 9】

Bit63

0

エラーコード	宛先ノード コード	送信元ノード コード	重 障 害 フ ラ グ	エラー情報
--------	--------------	---------------	----------------------------	-------

【書類名】 要約書

【課題】 複数のノードが該ノードを制御するサービスプロセッサに接続されてなるノード群が複数集まって形成された大規模プラットフォームを用いたマルチプロセッサシステムにおいて、障害処理を迅速かつ的確におこなうことができる障害処理方法等を提供する。

【課題手段】 障害ノードがサービスプロセッサに障害の発生を通知し、サービスプロセッサが障害情報を取得してサービスプロセッサマネージャに送信する工程と、障害ノードが同一のパーティションの他ノードに障害情報を伝達する工程と、同一のパーティションの他ノードがサービスプロセッサに障害の発生を通知し、サービスプロセッサが障害情報を取得してサービスプロセッサマネージャに送信する工程と、サービスプロセッサマネージャが障害情報に基づいてサービスプロセッサに指示し、サービスプロセッサが障害処理をおこなう工程とを含む。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 4 2 3 7]

1. 変更年月日	1 9 9 0 年 8 月 2 9 日
[変更理由]	新規登録
住 所	東京都港区芝五丁目 7 番 1 号
氏 名	日本電気株式会社